

$O(1)$ contribution of saddle point fluctuations to the free energy of Bethe Ansatz systems

F. Woynarovich*

*Institute for Solid State Physics and Optics
Hungarian Academy of Sciences
1525 Budapest 114, Pf 49.*

Abstract

We develop a method to calculate the contribution of the saddle-point fluctuations to the partition function of systems soluble by the Bethe Ansatz. Using this method we give the $O(1)$ corrections to the free energy of the 1D repulsive δ Bose gas both for periodic boundary conditions and for the open end case. We also generalize our method to more complicated systems and discuss the case of XXZ Heisenberg chain in more details.

PACS numbers: 02.30.Ik, 05.30.-d, 73.90.+f, 75.10.Pg, 75.10.Hr

Typeset using REVTEX

*e-mail: fw@szfki.hu

I. INTRODUCTION

Thermodynamic potentials such as the free energy are usually calculated in the infinite size limit, and in this limit only the leading (macroscopic) contribution is taken, as in studying bulk or macroscopic properties this gives a sufficient accuracy. Recent developments both in field theory and solid state physics have posed series of problems, where surface or impurity contributions are important [1,2]. As these contributions are less than the macroscopic ones, the otherwise formidable problem of calculating the next to leading order terms is set in focus.

For the one dimensional (1D) systems soluble by Bethe Ansatz (BA) the free energy is calculated following the method developed by C.N. Yang and C.P. Yang [3] for the δ Bose gas. The basic idea of this method is that through the density of the momenta (rapidities) an entropy can be defined enabling one to write up the free energy functional of the finite temperature system. Minimization of this functional with respect to the momentum density yields the most probable distribution of the momenta and the value of the (macroscopic) free energy. In terms of the partition function, the minimizing of the free energy functional corresponds to finding the state entering with highest weight into the partition function (saddle point). The contribution of the states around this one (saddle-point fluctuations) is an $O(1)$ factor to the partition function, i.e. an $O(1)$ correction in the free energy (if not zero). Our aim is to develop a convergent functional integral method to calculate this contribution, what is actually a finite size correction to the free energy.

First, to avoid special difficulties due to the BA, we write up, evaluate and discuss the convergence problems associated to the functional integral for the case of the *free Fermi gas*. Using the repulsive δ Bose gas – the simplest BA system – as an example, we generalize the procedure to the BA systems. We give also a general expression for the contribution of the saddle point fluctuations in more complicated systems, and discuss the case of XXZ Heisenberg chain in more detail.

The quantity we calculate for the free Fermi and δ Bose gas is the grand canonical partition function and the thermodynamical potential defined by its logarithm (grand canonical potential). In a strict sense this latter is different from the free energy, nevertheless it is not unusual to name it so. Throughout the paper also we use this certainly less precise, but hopefully not confusing name.

We consider systems at periodic boundary conditions (PBC). As in this case there is no boundary, the naïve conception, which identifies the $O(1)$ corrections as boundary contributions [1] would lead to the conclusion, that for PBC there are no $O(1)$ corrections. Surprisingly we find this is true only for the case of free Fermi gas, but does not hold for the interacting systems. In the case of the repulsive δ Bose gas we calculate the contribution of the saddle-point fluctuations also for open end systems. We find that also in this case the corrections are given by bulk properties (the energy of the particles as a function of the

rapidities and by the kernel, i.e. the derivative of the scattering phase shifts), and the nature of the boundary is reflected in the structure of the contributions.

We have to note, that our method works in the thermodynamic (infinite size) limit only, and as a consequence the correction we find gives the so-called residual entropy in the zero temperature limit [1]. For finite systems the entropy S tends to the logarithm of the ground-state degeneracy $\ln D$ as the temperature T tends to zero. The residual entropy we find behaves differently: its $T \rightarrow 0$ limit is not connected to the ground-state degeneracy D rather it depends on other properties like particle density or magnetization. This is a manifestation of the fact, that the size $\rightarrow \infty$ and $T \rightarrow 0$ limits do not commute.

The usual structure of the Bethe Ansatz equations makes it possible to give a general expression for the contribution of the saddle point fluctuations in a large class of Bethe Ansatz systems. Using the XXZ chain as an example we discuss some points one has to pay attention to when applying our formula. This considerations show that in addition to the saddle point fluctuations other effects can give $O(1)$ corrections too, more over in the isotropic ($SU(2)$ symmetric) case the degeneracy of the states belonging to the same spin-multiplet can give an even larger contribution.

The paper is organized as follows. In Sec. II we develop the method and discuss the convergence problems. We use the free Fermi gas to have a possibility to check the result. The method is generalized for the δ Bose gas in Sec. III. Here we treat both the PBC case and the case of open ends with surface potentials. Sec. IV is devoted to the generalization of the method for more complicated systems. Our results are summarized in Sec. V. In the bulk of the paper we concentrate on the main line of our method, and the technicalities are collected in appendices.

II. THE FREE FERMI GAS

The state of a free Fermion system is characterized by a set of wavenumbers

$$k_i = \frac{2\pi}{L} J_i. \quad (2.1)$$

where L is the size and at PBC the J quantum numbers are integers. The energy of a state is

$$E = \sum_i e(k_i) \quad \text{where} \quad e(k) = (k^2 - \mu) \quad (2.2)$$

with μ being the chemical potential, by which the required particle number N is fixed [4], and the grand canonical partition function is

$$Z = \sum_{\{k_i\}} \exp \left\{ -\beta \sum_i e(k_i) \right\}. \quad (2.3)$$

Here the summation over $\{k_i\}$ means summation over all possible k_i sets, and β is the inverse temperature $1/T$.

Now we follow the method developed by C.N. Yang and C.P. Yang [3] for the BA systems. We split up the k axis into intervals Δk , and introduce the density of ks ($\rho(k)$) and holes ($\rho_h(k)$) so that the number of the wavenumbers and holes in the $(k, k + \Delta k)$ interval is given by $L\Delta k\rho(k)$ and $L\Delta k\rho_h(k)$ respectively. Obviously

$$\rho(k) + \rho_h(k) = \frac{1}{2\pi}. \quad (2.4)$$

The number of states characterized by the same $\rho(k)$ function is

$$\Omega[\rho(k)] = \prod_k \omega(\rho(k)), \quad \text{with} \quad \omega(\rho(k)) = \binom{L\Delta k(\rho(k) + \rho_h(k))}{L\Delta k\rho(k)}, \quad (2.5)$$

where in $\prod_k k$ labels the Δk intervals. In the following $\ln \omega(\rho(k))$ is calculated by Stirling's formula

$$\ln \omega(\rho(k)) = L\Delta k s(\rho(k)) + \varsigma(\rho(k)) + O\left(\frac{1}{L\Delta k\rho(k)}, \frac{1}{L\Delta k\rho_h(k)}\right) \quad (2.6)$$

where

$$s(\rho(k)) = \{(\rho(k) + \rho_h(k)) \ln(\rho(k) + \rho_h(k)) - \rho(k) \ln \rho(k) - \rho_h(k) \ln \rho_h(k)\}, \quad (2.7)$$

and

$$\varsigma(\rho(k)) = -\frac{1}{2} \ln(2\pi L\Delta k) + \frac{1}{2} \ln \frac{\rho(k) + \rho_h(k)}{\rho(k) \cdot \rho_h(k)}. \quad (2.8)$$

By means of $\omega(\rho(k))$ a free energy functional is defined

$$F[\rho(k)] = \sum_k (Le(k)\rho(k)\Delta k - T \ln \omega(\rho(k))), \quad (2.9)$$

by which

$$Z = \sum_{\{\rho(k)\}} e^{-\beta F[\rho(k)]}. \quad (2.10)$$

(Here the summation extends over all possible $\rho(k)$ distributions.) In the usual procedure $\ln \omega(\rho(k))$ is calculated up to leading order in L only, i.e.

$$\ln \omega(\rho(k)) \cong L \Delta k s(\rho(k)) \quad (2.11)$$

is taken, and the free energy functional

$$F_L[\rho(k)] = L \sum_k (e(k)\rho(k) - Ts(\rho(k))) \Delta k \quad (2.12)$$

is minimized with respect to $\rho(k)$ leading to

$$\frac{\rho_0(k)}{\rho_{h,0}(k)} = e^{-\beta e(k)} \quad (2.13)$$

and

$$F_{\min} = -\frac{L}{\beta 2\pi} \sum \ln (1 + e^{-\beta e(k)}) \Delta k = -\frac{L}{\beta 2\pi} \int_{-\infty}^{\infty} \ln (1 + e^{-\beta e(k)}) dk. \quad (2.14)$$

Now we can define the functional integral for Z . We evaluate the summation over the possible $\rho(k)$ distributions through integrals over the $\rho(k)$ s: as $L\Delta k \rho(k)$ (being the number of ks in the interval Δk) is integer, the correct approximation for the summs is

$$\sum_{\{\rho(k)\}} \Rightarrow \int \cdots \int \prod_k (L\Delta k d\rho(k)) . \quad (2.15)$$

Introducing

$$r(k) = \rho(k) - \rho_0(k) \quad (2.16)$$

we can expand $F[\rho(k)]$ around $\rho_0(k)$. In order to have the functional integral convergent, in $\omega(\rho(k))$ of (2.6) we have to take into account the terms next to the leading ones too ($\varsigma(\rho)$). This gives

$$Z = e^{-\beta F_{\min}} \times \int \cdots \int \prod_k (L\Delta k dr(k)) \exp \left\{ - \sum_k \left(L\Delta k \frac{1}{2} \frac{\rho_0 + \rho_{h,0}}{\rho_0 \rho_{h,0}} (r(k))^2 - \frac{1}{2} \ln \frac{\rho_0 + \rho_{h,0}}{\rho_0 \rho_{h,0} L\Delta k 2\pi} \right) \right\} . \quad (2.17)$$

Here in the exponent the first term comes from the expansion of $F_L[\rho_0 + r]$, and the second term is $\varsigma(\rho_0)$. Introducing new variables

$$\xi(k) = \sqrt{L\Delta k \frac{1}{2} \frac{\rho_0(k) + \rho_{h,0}(k)}{\rho_0(k) \rho_{h,0}(k)}} r(k) \quad (2.18)$$

one arrives at

$$Ze^{\beta F_{\min}} = \prod_k \left(\frac{1}{\pi} \int e^{-(\xi(k))^2} d\xi(k) \right) = 1, \quad \text{i.e.} \quad Z = e^{-\beta F_{\min}}. \quad (2.19)$$

We can build up confidence in this calculation, as it reproduces the exact result. As, however, we want to generalize the method to cases where no exact results exist, it is worth to examine it in detail: it involves several approximations and also an implicit cut-off procedure, and we should see the conditions under which these are valid. Actually we have to check two types of approximations:

1. *The application of the free energy functional.* The question to be answered is whether the definition of the free energy functional is accurate enough. The partition functions (2.3) and (2.10) are certainly equivalent in leading order, but are not exactly equal, as (2.10) is obtained from (2.3) by a kind of averaging: in each Δk interval we take

$$\sum_{\{k_j\}} \exp \left\{ -\beta \sum_{k_j} e(k_j) \right\} \longrightarrow \exp \left\{ -\beta e(\bar{k}) L \rho(k) \Delta k + \ln \omega(\rho(k)) \right\}, \quad (2.20)$$

with $\sum_{\{k_j\}}$ meaning summation over all possible choices of $L \rho(k) \Delta k$ k_j s out of the $L(\rho(k) + \rho_h(k)) \Delta k$ possibilities, and \bar{k} being a k value within the interval Δk . As we are interested in calculating corrections to the macroscopic free energy we have to define this averaging more precisely. In Appendix A we show, that taking for \bar{k} the mean value of the k s in Δk we introduce an $O((\Delta k)^2)$ error to the free energy *density*, that disappears when the $\sum(\dots) \Delta k \longrightarrow \int(\dots) dk$ limit is taken. In a strict sense the mean value of the k s in Δk still depends on $\rho(k)$ and $\rho_h(k)$, but in the practice one may take the middle of Δk as \bar{k} . Although this introduces an $O(1/L)$ uncertainty in the mean value of energy/particle in each Δk interval, as however, these are random, do not sum up to an $O(1)$ contribution.

2. *The approximations applied while evaluating (2.10):*

- *Conditions for the application of Stirling's formula.* Each ξ integral collects the main part of the contribution from a region $-X < \xi < X$ where X is a number of the order of 5-6 ([5]). We may use Stirling's formula in the above form, if in the corresponding ρ region both $L \Delta k \rho(k)$, $L \Delta k \rho_h(k) \gg 1$. These lead to the requirements

$$L \Delta k \rho_0(k) \left(1 - X \sqrt{\frac{2 \rho_{h,0}(k)}{\rho_0(k) + \rho_{h,0}(k)} \frac{1}{L \Delta k \rho_0(k)}} \right) \gg 1 \quad (2.21)$$

and

$$L\Delta k \rho_{h,0}(k) \left(1 - X \sqrt{\frac{2\rho_0(k)}{\rho_0(k) + \rho_{h,0}(k)} \frac{1}{L\Delta k \rho_{h,0}(k)}} \right) \gg 1, \quad (2.22)$$

i.e. both condition are satisfied if

$$L\Delta k \rho_0(k) \gg 1 \quad \text{and} \quad L\Delta k \rho_{h,0}(k) \gg 1. \quad (2.23)$$

As for large enough k $\rho_0(k) \rightarrow 0$, one has to introduce a cutoff (say Λ) in the k space so that (2.23) are met for all $|k| < \Lambda$, i.e.

$$L\Delta k \rho_0(\Lambda) \gg 1. \quad (2.24)$$

This imposes a relation on the L and Λ . Similarly, if the temperature T is small enough $\rho_{h,0}(0) \rightarrow 0$ thus the method is applicable at large enough size but not at exactly zero temperature: the larger the size of the system is, the nearer the $T = 0$ can be approached.

We have to note, however, that since the entropy part of the *macroscopic* free energy functional is already obtained through Stirling's formula, the problem connected to the application of this approximation is in principle present in the calculation of the macroscopic part of the free energy too (Appendix D).

- *Limits of the ξ integrals.* The estimations (2.21-2.22) show also, that in case of (2.23) the limits of the ξ integrals can be taken to $\pm\infty$.
- *The $\varsigma(\rho_0)$.* We have taken $\varsigma(\rho)$ at ρ_0 and we neglected $\varsigma'(\rho_0)r + \varsigma''(\rho_0)r^2/2$. The term linear in r shifts the center of the ξ integral, while the quadratic term modifies the coefficient of ξ^2 . It is not hard to see that these corrections in the exponent are $O(1/L\Delta k \rho_0(k), 1/L\Delta k \rho_{h,0}(k))$ ones, what we may neglect if (2.23) holds.

Now we may conclude, that our functional integral method is established in a strict mathematical sense in the $L \rightarrow \infty$ limit only, and involves a cutoff procedure in the momentum space. We discuss the relation of the cutoff to the size and some details of the cutoff procedure (paying special attention to its connection to the macroscopic free energy) in Appendix D.

This method is not applicable for $T = 0$, but the $T \rightarrow 0$ limit is meaningful. (In the present case the $T \rightarrow 0$ limit reproduces the $T = 0$ result, but this is not so in the case of interacting systems, as we shall see later.)

III. THE REPULSIVE δ BOSE GAS

The case of periodic boundary conditions

The Hamiltonian defining the 1D repulsive δ Bose gas is

$$\hat{H} = - \sum_i \frac{\partial^2}{\partial x_i^2} + 2c \sum_{i < j} \delta(x_i - x_j), \quad c > 0. \quad (3.1)$$

Using the Bethe Ansatz Lieb and Liniger has shown, that the diagonalization of this Hamiltonian in a box with a length of period L can be reduced to the solution of the system of algebraic equations

$$Lk_i = 2\pi J_i - \sum_j^N 2 \tan^{-1} \left(\frac{k_i - k_j}{c} \right), \quad J_i = \frac{N-1}{2} \pmod{1}, \quad (3.2)$$

with N being the number of particles in the system [6]. The energy of the system is again given by (2.2), but now the wavenumbers are determined by the Eqs.(3.2) instead of Eq.(2.1). The finite temperature description [3] goes like we treated the free electron gas with the difference, that due to Eq.(3.2) Eq.(2.4) is replaced by

$$\begin{aligned} \rho(k) + \rho_h(k) &= \sigma + \sum_{k'} K(k, k') \rho(k') \Delta k' \\ &= \sigma + \int_{-\infty}^{\infty} K(k, k') \rho(k') dk' \end{aligned} \quad (3.3)$$

with

$$\sigma = \frac{1}{2\pi} \quad \text{and} \quad K(k, k') = \frac{1}{2\pi} \frac{2c}{c^2 + (k - k')^2}. \quad (3.4)$$

One should note, that only those $\rho(k)$ distributions are meaningful (physical), for which this equation yields $\rho_h(k) \geq 0$ for all k . The free energy is again given by (2.12), but now it is constrained by (3.3). Its minimization leads to

$$\frac{\rho_0(k)}{\rho_{h,0}(k)} = e^{-\beta\epsilon(k)}, \quad (3.5)$$

with the energy $\epsilon(k)$ determined by the equation

$$\begin{aligned} \epsilon(k) &= e(k) - T \sum_{k'} K(k, k') \ln \left(1 + e^{-\beta\epsilon(k')} \right) \Delta k' \\ &= e(k) - T \int_{-\infty}^{\infty} K(k, k') \ln \left(1 + e^{-\beta\epsilon(k')} \right) dk'. \end{aligned} \quad (3.6)$$

Once $\epsilon(k)$ is found, $\rho_0(k)$ is given by the equation

$$\rho_0(k) = \frac{1}{1 + e^{\beta\epsilon(k)}} \sigma + \int_{-\infty}^{\infty} \frac{1}{1 + e^{\beta\epsilon(k)}} K(k, k') \rho_0(k') dk', \quad (3.7)$$

and the minimal free energy is

$$F_{\min} = -\frac{L}{\beta 2\pi} \sum_k \ln \left(1 + e^{-\beta \epsilon(k)} \right) \Delta k = -\frac{L}{\beta 2\pi} \int_{-\infty}^{\infty} \ln \left(1 + e^{-\beta \epsilon(k)} \right) dk. \quad (3.8)$$

To continue we expand the (2.12) around $\rho_0(k)$ and $\rho_{h,0}(k)$:

$$F_L [\rho(k)] \simeq F_{\min} - T \sum_k L \Delta k \frac{1}{2} \left(\frac{(r(k) + r_h(k))^2}{\rho_0(k) + \rho_{h,0}(k)} - \frac{r_h^2(k)}{\rho_{h,0}(k)} - \frac{r^2(k)}{\rho_0(k)} \right). \quad (3.9)$$

Here the quantities

$$r(k) = \rho(k) - \rho_0(k) \quad \text{and} \quad r_h(k) = \rho_h(k) - \rho_{h,0}(k) \quad (3.10)$$

are constrained due to (3.3) by

$$r(k) + r_h(k) = \sum_{k'} K(k, k') r(k') \Delta k'. \quad (3.11)$$

This way the functional integral for the partition function reads

$$Z = e^{-\beta F_{\min}} \int \cdots \int \prod_k (L \Delta k dr(k)) \times \exp \left\{ - \sum_k \left(L \Delta k \frac{1}{2} \left(\frac{(r(k) + r_h(k))^2}{\rho_0(k) + \rho_{h,0}(k)} - \frac{r_h^2(k)}{\rho_{h,0}(k)} - \frac{r^2(k)}{\rho_0(k)} \right) - \frac{1}{2} \ln \frac{\rho_0 + \rho_{h,0}}{\rho_0 \rho_{h,0} L \Delta k 2\pi} \right) \right\}, \quad (3.12)$$

what is to be evaluated under the constrain (3.11). After eliminating $r_h(k)$ we arrive at

$$\begin{aligned} & \sum_k L \Delta k \frac{1}{2} \left(\frac{(r(k) + r_h(k))^2}{\rho_0(k) + \rho_{h,0}(k)} - \frac{r_h^2(k)}{\rho_{h,0}(k)} - \frac{r^2(k)}{\rho_0(k)} \right) = \\ & -\frac{1}{2} \sum_{k, k', k''} L \Delta k r(k) K(k', k) \frac{\rho_0(k') \Delta k'}{\rho_{h,0}(k') (\rho_0(k') + \rho_{h,0}(k'))} K(k', k'') r(k'') \Delta k'' \\ & + \sum_{k, k'} L \Delta k r(k) \frac{K(k, k')}{\rho_{h,0}(k)} r(k') \Delta k' \\ & - \frac{1}{2} \sum_k L \Delta k r^2(k) \frac{\rho_0(k) + \rho_{h,0}(k)}{\rho_0(k) \rho_{h,0}(k)}. \end{aligned} \quad (3.13)$$

Now, just as in the case of free Fermi gas we introduce $\xi(k)$ according to Eq.(2.18). This leads to

$$- \sum_{k, k', k''} \xi(k) (\delta_{k', k} - M_{k', k}) (\delta_{k', k''} - M_{k', k''}) \xi(k''), \quad (3.14)$$

where $\xi(k)$ is that of (2.18), $\delta_{k, k'}$ is the Kronecker symbol, and

$$M_{k, k'} = \frac{1}{\rho_{h,0}(k)} \sqrt{\frac{\rho_0(k)\rho_{h,0}(k)\Delta k}{\rho_0(k) + \rho_{h,0}(k)}} K(k, k') \sqrt{\frac{\rho_0(k')\rho_{h,0}(k')\Delta k'}{\rho_0(k') + \rho_{h,0}(k')}} \quad (3.15)$$

Finally, changing the integration variable to ξ in (3.12) we have

$$\begin{aligned} Z &= e^{-\beta F_{\min}} \int \cdots \int \prod_k \left(\frac{d\xi(k)}{\pi} \right) \exp \left\{ - \sum_{k, k', k''} \xi(k) (\delta_{k, k'} - M_{k, k'}) (\delta_{k', k''} - M_{k', k''}) \xi(k'') \right\} \\ &= e^{-\beta F_{\min}} (\det [\delta_{k, k'} - K_{k, k'}])^{-1} \end{aligned} \quad (3.16)$$

with

$$K_{k, k'} = \sqrt{\rho_{h,0}(k)} \cdot M_{k, k'} \cdot \frac{1}{\sqrt{\rho_{h,0}(k')}} = \sqrt{\frac{\rho_0(k)\Delta k}{\rho_0(k) + \rho_{h,0}(k)}} K(k, k') \sqrt{\frac{\rho_0(k')\Delta k'}{\rho_0(k') + \rho_{h,0}(k')}}. \quad (3.17)$$

The determinant in the above formula evaluated using the identities

$$(\det [\mathbf{I} - \mathbf{K}])^{-1} = \exp \{ -\text{Tr} \ln(\mathbf{I} - \mathbf{K}) \} = \exp \left\{ \sum_n \frac{1}{n} \text{Tr} \mathbf{K}^n \right\} \quad (3.18)$$

and taking the $\sum_k \Delta k \rightarrow \int dk$ limit yields

$$Z = e^{-\beta F_{\min} + \Delta \mathcal{S}} \quad (3.19)$$

where

$$\Delta \mathcal{S} = \sum_n \frac{1}{n} \mathcal{K}^n \quad (3.20)$$

with

$$\begin{aligned} \mathcal{K}^n &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dk_1 \cdots dk_n \times \\ &\quad \frac{\rho_0(k_1)}{\rho_0(k_1) + \rho_{h,0}(k_1)} K(k_1, k_2) \frac{\rho_0(k_2)}{\rho_0(k_2) + \rho_{h,0}(k_2)} \cdots \frac{\rho_0(k_n)}{\rho_0(k_n) + \rho_{h,0}(k_n)} K(k_n, k_1) \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dk_1 \cdots dk_n \frac{1}{1 + e^{\beta \epsilon(k_1)}} K(k_1, k_2) \frac{1}{1 + e^{\beta \epsilon(k_2)}} \cdots \frac{1}{1 + e^{\beta \epsilon(k_n)}} K(k_n, k_1). \end{aligned} \quad (3.21)$$

(We have to note here, that, although our notation suggests so, for $T \neq 0$ $\Delta\mathcal{S}$ is not a correction to the entropy, nevertheless we use this notation, as formally it comes from the density of states.)

Now, similarly to the case of the free Fermi gas we should make a kind of "validity test". In this, in addition to the questions discussed there (application of the free energy functional, and approximations in the evaluation of the partition function) one has to see also, that the $\sum_n \frac{1}{n} \mathcal{K}^n$ is convergent.

1. *The free energy functional.* While in the case of the free Fermi gas the application of the free energy functional introduced to the free energy density an $O((\Delta k)^2)$ error only, in the present case we are faced to an apparently more serious problem, which originates from the fact, that the J_i quantum numbers in (3.2) are either integers or half-integers depending on the parity of the particle number. Changing all J_i quantum numbers from integer to half-odd-integers or vice versa shifts all k_i by π/L leading to a shift also in the free energy. For general $\rho(k)$ s this shift can be of $O(1)$ suggesting that the free energy for this system as a function of the $\rho(k)$ is defined with an $O(1)$ accuracy only. Fortunately not this is the case. In Appendix B we show, that the above used definition of the free energy is accurate enough for all the $\rho(k)$ s contributing to the Z significantly, as the uncertainty coming from the prescriptions for the J_i s is $O(1/\sqrt{L})$, i.e. it disappears in the $L \rightarrow \infty$ limit.
2. *The approximations in evaluating the Z .* This group of questions is completely analogous to the questions emerged in connection to the free Fermi gas, and the answers are similar too: in an appropriate cutoff procedure involving the limits $L \rightarrow \infty$, $\Delta k \rightarrow 0$ and $\Lambda \rightarrow \infty$ (Appendix D) Stirling's formula can be applied in its (2.6) form, $\varsigma(\rho)$ can be taken at ρ_0 and $\rho_{h,0}$ and also the limits of the ξ integrals can be taken to infinity. (The major point in this is that the variable $r \propto \xi/\sqrt{L\Delta k}$, i.e. only a $\sim 1/\sqrt{L\Delta k}$ neighborhood of the ρ_0 and $\rho_{h,0}$ plays any role.)
3. *The application of the formulas (3.18)* (the convergence of (3.20)). The condition for this is that all the eigenvalues of the matrix \mathbf{K} are of modulus less than one. In the Appendix C we show that this is true for any $T > 0$, and we show also, that the $T \rightarrow 0$ limit of the sum $\sum_n \frac{1}{n} \mathcal{K}^n$ exists too.

As we have seen, our method is strictly established in the $L \rightarrow \infty$ limit only, but in a less strict manner we may say, however, that the $O(1)$ corrections are correctly given also for finite but large enough size too: although in that case the $\Lambda \rightarrow \infty$ limit can not be completed (as the $L\Delta k\rho(\Lambda) \gg 1$ would not hold for very large Λ), the contribution of the states above Λ would be suppressed anyhow due to the large energy. For large but finite system, however, one has to be careful with the $T \rightarrow 0$ limit. The quantity $\Delta\mathcal{S}$ is an $O(1)$ correction to the thermodynamic potential:

$$-T \ln Z = F_{\min} - T\Delta\mathcal{S}. \quad (3.22)$$

As for $T \rightarrow 0$ $F_{\min} \rightarrow E_0$ (with E_0 being the ground state energy) $\lim_{T \rightarrow 0} \Delta\mathcal{S}$ is an entropy. For infinite L this is the residual entropy which can be finite, but for large but finite L $\lim_{T \rightarrow 0} \Delta\mathcal{S}$ should be zero as the ground state is non degenerated. The resolution of this contradiction is that for finite L our calculation breaks down at temperatures where $L\Delta k\rho_{0,h}(0) \sim 1$, and below this temperatures the $O(1)$ corrections gradually disappear. For this reason the $L \rightarrow \infty$ and $L \gg 1$ cases should be distinguished!

The case of open ends

The system, just as in the previously discussed case, is described by the Hamiltonian (3.1), but the quantization condition is different: now the ring is not closed, and the particles are reflected on the ends. We suppose, that if a particle with wavenumber $-k$ arrives at the end at $x = 0$, it is reflected to have a wavenumber k while the phase of the wavefunction is shifted by $\varphi_0(k)$:

$$e^{-ikx} \longrightarrow e^{ikx+i\varphi_0(k)} \quad (x \sim 0). \quad (3.23)$$

Similarly, a particle of wavenumber k arriving at the end at $x = L$ is reflected to have a wavenumber $-k$ while the phase of the wavefunction is shifted by $\varphi_L(k)$:

$$e^{ikx} \longrightarrow e^{-ikx+2ikL+i\varphi_L(k)} \quad (x \sim L). \quad (3.24)$$

(If the system is closed by infinitely high potential walls, $\varphi_0(k) = \varphi_L(k) = \pi$, but with other choices of $\varphi_0(k)$ and $\varphi_L(k)$ different types of ends can be generated. In Appendix E we discuss a case, when a surface potential having bound states closes the chain.) The Bethe Ansatz equations of such a system read:

$$2Lk_i = 2\pi J_i - \varphi_0(k_i) - \varphi_L(k_i) - \sum_{j(\neq i)}^N \left\{ 2 \tan^{-1} \left(\frac{k_i - k_j}{c} \right) + 2 \tan^{-1} \left(\frac{k_i + k_j}{c} \right) \right\}. \quad (3.25)$$

Here the J_i numbers are integers, $k_j \neq \pm k_i$ for $j \neq i$, and none of the k_j s equals zero. This means, the real k_j are different positive numbers. Now we suppose, the $\varphi_0(k)$ and $\varphi_L(k)$ phases do not generate surface bound states (i.e. states with complex ks) and we discuss the distribution of the real ks (the generalization for the case of surface bound states is straightforward, as it is seen in Appendix E).

To proceed we split up the positive k axis into Δk intervals, and introduce the densities of the particles and holes ($\bar{\rho}(k)$ resp. $\bar{\rho}_h(k)$) in the usual manner. Now the integral equation connecting these quantities is

$$\begin{aligned} \bar{\rho}(k) + \bar{\rho}_h(k) &= \sigma(k) + \sum_{k'} \bar{K}(k, k') \bar{\rho}(k') \Delta k' \\ &= \sigma(k) + \int_0^\infty \bar{K}(k, k') \bar{\rho}(k') dk' \end{aligned} \quad (3.26)$$

with

$$\sigma(k) = \frac{1}{2\pi} \left\{ 2 + \frac{1}{L} \frac{\partial \varphi_0(k)}{\partial k} + \frac{1}{L} \frac{\partial \varphi_L(k)}{\partial k} - \frac{1}{L} \frac{4c}{c^2 + (2k)^2} \right\}, \quad (3.27)$$

and

$$\overline{K}(k, k') = \frac{1}{2\pi} \left(\frac{2c}{c^2 + (k - k')^2} + \frac{2c}{c^2 + (k + k')^2} \right). \quad (3.28)$$

Now the number of states characterized by the same $\overline{\rho}(k)$ function is

$$\overline{\Omega}[\overline{\rho}(k)] = \left(\frac{L\Delta k(\overline{\rho}(k) + \overline{\rho}_h(k)) - 1/2}{L\Delta k\overline{\rho}(k)} \right) \Big|_{k=0} \prod_{k \neq 0} \omega(\overline{\rho}(k)). \quad (3.29)$$

Here $k = 0$ refers to the interval beginning at the origin, and

$$\omega(\overline{\rho}(k)) = \left(\frac{L\Delta k(\overline{\rho}(k) + \overline{\rho}_h(k))}{L\Delta k\overline{\rho}(k)} \right), \quad (3.30)$$

just as previously. The contribution of the $k = 0$ interval differs from those of the others as the $k = 0$ wavenumber (what is right on the edge of the interval) is not allowed. (In a $2\Delta k$ interval containing the origin in the middle there are $2L\Delta k(\overline{\rho}(k = 0) + \overline{\rho}_h(k = 0))$ k values, out of this $L\Delta k(\overline{\rho}(k) + \overline{\rho}_h(k)) - 1/2$ is positive, thus this is the number of choosable ks in the $k = 0$ interval.) Applying Stirling's formula we arrive at

$$\overline{\Omega}[\overline{\rho}(k)] = \sqrt{\left(\frac{\overline{\rho}_h(k)}{\overline{\rho}(k) + \overline{\rho}_h(k)} \right)} \Big|_{k=0} \prod_k \omega(\overline{\rho}(k)). \quad (3.31)$$

The free energy functional is

$$\overline{F}[\overline{\rho}(k)] = \sum_k (Le(k)\overline{\rho}(k)\Delta k - T \ln \omega(\overline{\rho}(k))) - \frac{T}{2} \ln \frac{\overline{\rho}_h(k)}{\overline{\rho}(k) + \overline{\rho}_h(k)} \Big|_{k=0}. \quad (3.32)$$

The minimization of the free energy leads to the condition

$$\frac{\overline{\rho}_0(k)}{\overline{\rho}_{h,0}(k)} = e^{-\beta\epsilon(k)}, \quad (3.33)$$

with $\epsilon(k)$ given by (3.6) (just as in the PBC case), and it yields a minimal value

$$\begin{aligned} \overline{F}_{\min} &= -\frac{L}{\beta} \sum \ln(1 + e^{-\beta\epsilon(k)}) \sigma(k)\Delta k + \frac{T}{2} \ln(1 + e^{-\beta\epsilon(0)}) \\ &= -\frac{L}{\beta} \int_0^\infty \ln(1 + e^{-\beta\epsilon(k)}) \sigma(k)dk + \frac{T}{2} \ln(1 + e^{-\beta\epsilon(0)}). \end{aligned} \quad (3.34)$$

Finally the evaluation of the functional integral leads to

$$\overline{Z} = e^{-\beta \overline{F}_{\min} + \overline{\Delta S}} \quad (3.35)$$

with

$$\overline{\Delta S} = \sum_n \frac{1}{n} \overline{\mathcal{K}}^n \quad (3.36)$$

where

$$\begin{aligned} \overline{\mathcal{K}}^n &= \int_0^\infty \cdots \int_0^\infty dk_1 \cdots dk_n \times & (3.37) \\ &\frac{\overline{\rho}_0(k_1)}{\overline{\rho}_0(k_1) + \overline{\rho}_{h,0}(k_1)} \overline{K}(k_1, k_2) \frac{\overline{\rho}_0(k_2)}{\overline{\rho}_0(k_2) + \overline{\rho}_{h,0}(k_2)} \cdots \frac{\overline{\rho}_0(k_n)}{\overline{\rho}_0(k_n) + \overline{\rho}_{h,0}(k_n)} \overline{K}(k_n, k_1) \\ &= \int_0^\infty \cdots \int_0^\infty dk_1 \cdots dk_n \frac{1}{1 + e^{\beta \epsilon(k_1)}} \overline{K}(k_1, k_2) \frac{1}{1 + e^{\beta \epsilon(k_2)}} \cdots \frac{1}{1 + e^{\beta \epsilon(k_n)}} \overline{K}(k_n, k_1). \end{aligned}$$

Using the actual form of $\sigma(k)$ and $\overline{K}(k, k')$ it is easy to see, that

$$\overline{F}_{\min} = F_{\min} + \Delta F + \phi_0 + \phi_L, \quad (3.38)$$

with

$$\Delta F = \frac{T}{2} \ln \left(1 + e^{-\beta \epsilon(0)} \right) + \frac{T}{2\pi} \int_0^\infty \frac{4c}{c^2 + (2k)^2} \ln \left(1 + e^{-\beta \epsilon(k)} \right) dk, \quad (3.39)$$

$$\phi_{0/L} = -\frac{T}{2\pi} \int_0^\infty \frac{\partial \varphi_{0/L}(k)}{\partial k} \ln \left(1 + e^{-\beta \epsilon(k)} \right) dk. \quad (3.40)$$

We note here the following. The macroscopic part of the thermodynamic potential F_{\min} is independent of the boundary condition as it should be. ΔF is a consequence of the openness of the chain, but the nature of the surfaces is reflected in ϕ_0 resp. ϕ_L only. These three $O(1)$ corrections directly connected to the surfaces are given already by the usual thermodynamic treatment. The contribution of the saddle point fluctuations depends on the boundary conditions in its structure. To see the difference between the ΔS valid for the PBC and $\overline{\Delta S}$ applying for the open ends we introduce the function

$$\widetilde{K}(k, k') = \sum_{n=1}^{\infty} \frac{1}{n} \widetilde{\mathcal{K}}^n(k, k') \quad (3.41)$$

with

$$\begin{aligned} \widetilde{K}^n(k, k') &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dk_1 \cdots dk_{n-1} \times \\ &\sqrt{\frac{1}{1 + e^{\beta\epsilon(k)}}} K(k, k_1) \frac{1}{1 + e^{\beta\epsilon(k_1)}} K(k_1, k_2) \frac{1}{1 + e^{\beta\epsilon(k_2)}} \cdots \frac{1}{1 + e^{\beta\epsilon(k_{n-1})}} K(k_{n-1}, k') \sqrt{\frac{1}{1 + e^{\beta\epsilon(k')}}}. \end{aligned} \quad (3.42)$$

Using the form of $K(k, k')$ and $\overline{K}(k, k')$ it is not hard to see, that with this notation

$$\mathcal{K}^n = \int_{-\infty}^{\infty} \widetilde{K}^n(k, k) dk, \quad \text{and} \quad \overline{\mathcal{K}}^n = \int_0^{\infty} \widetilde{K}^n(k, k) dk + \int_0^{\infty} \widetilde{K}^n(k, -k) dk, \quad (3.43)$$

thus

$$\Delta\mathcal{S} = \int_{-\infty}^{\infty} \widetilde{K}(k, k) dk, \quad \text{while} \quad \overline{\Delta\mathcal{S}} = \int_0^{\infty} \widetilde{K}(k, k) dk + \int_0^{\infty} \widetilde{K}(k, -k) dk. \quad (3.44)$$

We note that the considerations presented in Appendix C concerning the convergence of $\Delta\mathcal{S}$ hold for $\overline{\Delta\mathcal{S}}$ too, i.e. also $\overline{\Delta\mathcal{S}}$ together with its $T \rightarrow 0$ limit exists.

IV. GENERALIZATION

The repulsive δ Bose gas is the simplest BA system in the sense, that the particles have no internal structure and do not form bound states allowing to describe the system by one set of real parameters (the wavenumbers). In most of the BA systems, however, for the thermodynamic description one has to introduce many (in most of the cases *infinitely* many) sets of rapidities. These can be of different type. For example in the case of Heisenberg chain these variables are the centers of the strings of different length [9], for the repulsive δ Fermi gas one set gives the wavenumbers, and the others are strings connected to the spin state of the system, while in case of the Hubbard model there are three type of rapidities: the real wavenumbers, the centers of the strings connected to the spins and the centers of the strings connected to the bound pairs [10]. (Here we take granted, that the string hypothesis works, as all in the known cases, where independent check is possible, it gives the correct result [11].) These systems are much more complicated than the δ Bose gas, nevertheless it seems, that our calculation is generalizable for a larger class of them. Now we outline this procedure.

For the sake of simplicity we denote all of the rapidities by k , this will cause no confusion. To each set of rapidities particle and hole densities ($\rho^{(n)}(k)$ resp. $\rho_h^{(n)}(k)$) can be defined. Consider a system in which these satisfy a set of integral equations of the type:

$$\rho^{(n)}(k) + \rho_h^{(n)}(k) = \sigma_n(k) + \sum_m \int K_{n,m}(k, k') \rho^{(m)}(k') dk'. \quad (4.1)$$

The number of states described by the same set of densities is now supposed to be

$$\prod_n \Omega \left[\rho^{(n)}(k) \right] \quad (4.2)$$

and the free energy functional to be minimized is of the form

$$\mathbf{F} \left[\rho^{(n)}(k) \right] = L \sum_n \int \left(e_n(k) \rho^{(n)}(k) - T s \left(\rho^{(n)}(k) \right) \right) dk \quad (4.3)$$

where $e_n(k)$ is the energy of an object of type n and rapidity k . The minimization leads to the condition

$$\frac{\rho_0^{(n)}(k)}{\rho_{h,0}^{(n)}(k)} = e^{-\beta \epsilon_n(k)} \quad (4.4)$$

with

$$\epsilon_n(k) = e_n(k) - T \sum_m \int \ln \left(1 + e^{-\beta \epsilon_m(k')} \right) K_{m,n}(k', k) dk' . \quad (4.5)$$

The minimal value of the free energy functional is

$$\mathbf{F}_{\min} = -TL \sum_n \int \ln \left(1 + e^{-\beta \epsilon_n(k)} \right) \sigma_n(k) dk . \quad (4.6)$$

The contribution of the states near to the one minimizing the free energy functional can be calculated by the functional integral method described in the previous section. Through a very straightforward calculation one finds, that the contribution of these states to the free energy given as the logarithm of the partition function is

$$\Delta \mathbf{F} = -T \Delta \mathbf{S} , \quad (4.7)$$

with

$$\Delta \mathbf{S} = \sum_n \frac{1}{n} K^n , \quad (4.8)$$

where now

$$K^n = \sum_{m_1} \cdots \sum_{m_n} \int \cdots \int dk_1 \cdots dk_n \times \quad (4.9)$$

$$\frac{1}{1 + e^{\beta \epsilon_{m_1}(k_1)}} K_{m_1 m_2}(k_1, k_2) \frac{1}{1 + e^{\beta \epsilon_{m_2}(k_2)}} \cdots \frac{1}{1 + e^{\beta \epsilon_{m_n}(k_n)}} K_{m_n m_1}(k_n, k_1) .$$

The above calculation is a formal generalization of the procedure applied for the δ Bose gas, and it involves the same kinds of approximations too. For this in any case one should check,

if the conditions are met. Here we emphasize one point: Stirling's formula is applicable only if for all $\rho^{(n)}(k)$ giving significant contribution in the integral $L\rho^{(n)}(k)\Delta k \gg 1$ and $L\rho_h^{(n)}(k)\Delta k \gg 1$. This can be made true taking the $L \rightarrow \infty$ limit only in the case, if the $\rho_0^{(n)}(k) > 0$ and $\rho_{h,0}^{(n)}(k) > 0$ for all n and k .

Our result has a meaning only if the sum defining $\Delta\mathbf{S}$ converges. In any known system the kernel $K_{mm'}(k, k')$ is rather complicated, and checking the convergence may encounter difficulties. In some simple cases, however, this check is doable: For example in the case of the Heisenberg chain for $T = 0$ the density of 1-strings (real rapidities) remains finite only, all other densities disappear, and the procedure of Appendix C can be applied. This gives the result that in the case of finite magnetic field the sum is convergent for the complete antiferromagnetic region (the anisotropy $\varrho > -1$ in the Hamiltonian (4.10)), but it is not convergent in the critical region ($1 \geq \varrho > -1$) for zero field. This renders it likely, that the sum is convergent for finite temperature too for finite field for any ϱ , or even for zero field if $\varrho > 1$, but does not support any guess for finite temperature and no field if $1 \geq \varrho > -1$.

Even if the $\Delta\mathbf{S}$ exists, to decide if (4.7) is correct for a given system *may need further considerations*. In deriving (4.7) we assumed that only (4.1) constrains the densities and that in principle all of the states are so described, i.e. (4.2) gives correctly the number of states described by the same density. These assumptions, however, may not be true even in the best known cases. Two kinds of problems may arise:

1. In addition to (4.1) there are *other constraints* on the densities too. For example in systems of 1/2 spins the total number of the rapidities must not exceed the half of the number of sites/particles. (This expresses the fact, that the BA equations describe the states of total spin $S^z \geq 0$ only, the $S^z < 0$ states are obtained by reflection)
2. Not all of the states are automatically described by the possible densities. For example for SU(2) systems the BA equations describe the states of highest weight only, the others are obtained from these by further manipulations. It is also possible, that in the system there are more vacua, and more sets of excitations, and this multiplicity is not taken into account neither in the usual, nor in the present description.

In the following we discuss these problems briefly in case of the Heisenberg chain.

The XXZ Heisenberg chain of length N in magnetic field h is described by the Hamiltonian

$$\hat{H}_{XXZ} = \sum_i^N \left(S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + \varrho S_i^z S_{i+1}^z - h S_i^z \right). \quad (4.10)$$

This model is Bethe Ansatz diagonalisable, but the BA equations give certain classes of the eigenstates only, and the others should be constructed by further manipulations. As the basic properties of the model are concerned four different case should be distinguished,

but from the thermodynamic point of view the following three *antiferromagnetic* cases are important.

The $-1 < \varrho < 1$ planar case. In this case the BA equations give the $S^z = \sum_i S_i^z \geq 0$ states only: in all of the solutions the total number of turned down spins (spin-waves) is less than or equal to the half of the number of sites $N/2$. As it has been mentioned in point 1. above this imposes a constrain on the densities as they should satisfy the sum rule

$$\frac{1}{2} \geq \sum_{(n)} l_{(n)} \int \rho^{(n)}(k) dk, \quad (4.11)$$

with $l_{(n)}$ being the length of the string-type labeled by (n) . Those densities not obeying this constrain are nonphysical, this can be manifested in negative $\rho_h(k)$'s generated by (4.1). If in (4.11) the equality holds, than $S^z = 0$, if the $>$ sign is valid, than $S^z > 0$, and the $S^z < 0$ states are constructed by reflection of the $S^z > 0$ ones. This means, that calculating the partition function the contributions of the $S^z > 0$ states have to be taken into account with weights

$$1 + \exp\{-2\beta S^z h\}, \quad (4.12)$$

i.e. all $S^z > 0$ states have additional contributions to the free energy

$$-T \ln (1 + \exp\{-2\beta S^z h\}), \quad (4.13)$$

but no such contributions exist for the $S^z = 0$ states. For finite h the magnetization is macroscopic, $S^z \propto L$, thus if $L \rightarrow \infty$, the above contribution to the free energy disappears, but for $h = 0$ it remains finite $-T \ln 2$. On the other hand in finite magnetic field all the densities contributing significantly to the functional integral obey the sum rule (the equilibrium densities satisfy (4.11) with the sign $>$), thus the correction (4.7) is certainly correct, provided it is convergent. This is also true for zero magnetic field, although the situation in that case is somewhat different. In zero field the equilibrium densities describe states with zero magnetization (i.e. they satisfy (4.11) with the sign $=$), and the functional integral also involves nonphysical densities (which violate the sum rule). Taking into account the symmetry of the functional integral it is clear, however, that the contribution of the nonphysical densities is the same as the contribution of the densities describing $S^z > 0$ (physical) states, i.e. the functional integral takes into account the $S > 0$ states with a weight of two – just as it should be.

The $\varrho = 1$ isotropic chain. In this case in addition to the above constrain an other problem arises. The model has an SU(2) symmetry, and the BA equations describe the highest weight ($S^2 = S^z(S^z + 1)$) states only, the other states of the same spin length are obtained by means of the S^- operator. For this, if $h \neq 0$, the contribution to the partition function of one highest weight state should be weighted by

$$\frac{1 - \exp\{-\beta h(2S^z + 1)\}}{1 - \exp\{-\beta h\}} \quad (4.14)$$

to get the contribution of the complete multiplet, i.e. to the free energy functional a term

$$-T \ln \left(\frac{1 - \exp\{-\beta h(2S^z + 1)\}}{1 - \exp\{-\beta h\}} \right) \quad (4.15)$$

should be added. In the thermodynamic limit $\exp\{-\beta h(2S^z + 1)\}$ disappears, and we find an additional $O(1)$ correction

$$T \ln (1 - \exp\{-\beta h\}) \quad (4.16)$$

to the free energy. Note, that this is divergent if $h \rightarrow 0$ indicating, that the corresponding correction in the $h = 0$ case is of a different order of magnitude. (This divergence is a consequence of the thermodynamic limit, for finite N this term would behave as $-T \ln(2S^z(h) + 1)$.)

In the case of zero magnetic field all members of a spin multiplet are of the same energy, thus in order to get the contribution of a complete multiplet characterized by a spin length \mathcal{L} (this \mathcal{L} being equal to the S^z of the highest weight member) a correction

$$-T \ln(2\mathcal{L} + 1) \quad (4.17)$$

should be added to the free energy functional. Although for the equilibrium density this correction is zero, we show in Appendix F, that it can be very large when deviating from the equilibrium density, thus its treatment requires a much more subtle procedure, what is beyond the grasp of the present work. We also give an estimation according to which this term is $O(\ln N)$, so we have to conclude that for the isotropic chain in zero field the leading correction to the macroscopic free energy may come from this term, not from the saddle point fluctuations.

The $\varrho > 1$ "easy axis" anisotropic chain. In this case – similarly to the $\varrho < 1$ case – the solutions of the BA equations give all of the $S^z \geq 0$ states, and the $S^z < 0$ states are obtained by reflecting the $S^z > 0$ ones. In this sense the $\varrho > 1$ and $\varrho < 1$ are in close analogy and the conclusions concerning the $O(1)$ corrections made for $\varrho < 1$ hold also for $\varrho > 1$. Nevertheless in this case there is an additional problem. In this region the vacuum is twofold degenerated, this is manifested in the fact, that from the ground-state density of rapidities two different solution of the BA equations can be reconstructed. Also the low energy excited states can be grouped into two sets as being the excitations above one or the other vacuum [12], but it has not been studied yet, if such a degeneracy exists also in the thermodynamically important highly excited states. It is also not clear yet, if our method takes this kind of degeneracy automatically into account, although this can be important, as it is expected to give an $O(1)$ correction too, which behaves like $-T \ln 2$ as $T \rightarrow 0$.

V. SUMMARY

In the present work based on the method devised by Yang and Yang [3] we developed a functional integral method to calculate $O(1)$ corrections to the free energy of macroscopic BA integrable systems. In the Yang and Yang method the free energy of a system is written up as a functional of the momentum density, and this functional is minimized in order to find the actual value of the free energy. In terms of the grand canonical partition function the equilibrium density of momenta (at which the free energy is minimal) defines the states entering into the partition function with highest weights. The basic point of our calculation is that in evaluating the grand canonical partition function after the minimization of the free energy functional, (what actually gives the macroscopic part,) the contribution of the states near to the equilibrium (saddle point fluctuations) can be calculated by a Gaussian integral. To define this integral properly one needs to calculate the entropy entering into the free energy functional up to next to leading order.

In addition to the technical problems the calculation of non macroscopic corrections to the macroscopic free energy rises some conceptional questions too. The Yang and Yang method has been developed to pick up the leading contribution only, thus in calculating further terms one has to see, that this refinement is meaningful, the method is accurate enough to calculate the next to leading contributions too. This involves two kinds of problems. The first is if it is possible at all to define an accurate enough free energy density in terms of the momentum density. (Questions of this type are discussed in Appendices A and B.) The other kind of problem is connected to the accuracy by which the macroscopic part itself is calculated: in systems, in which the number of available momenta is infinite, a cutoff procedure must be introduced even to calculate the macroscopic free energy. (The problem of this cutoff procedure and its resolution in the thermodynamic limit is discussed in Appendix D.)

In order to avoid difficulties of special models we write up and discuss the method using the free Fermi gas. Our calculation reproduces the exact result, but we see that our method for the above mentioned implicit cutoff procedure is established in a strict mathematical sense in the thermodynamic limit only.

Next the method is generalized for the repulsive δ Bose gas with PBS. The structure of the BA equations for the rapidity densities of this system is (3.3)

$$\rho(k) + \rho_h(k) = \sigma(k) + \int_{-\infty}^{\infty} K(k, k') \rho(k') dk' \quad (5.1)$$

with

$$\sigma(k) = \frac{1}{2\pi}, \quad K(k, k') = \frac{1}{2\pi} \frac{2c}{c^2 + (k - k')^2}, \quad (5.2)$$

and the energy associated with a particle is

$$e(k) = k^2 - \mu. \quad (5.3)$$

The $O(1)$ correction to the free energy we find is of the form $-T\Delta\mathcal{S}$ with $\Delta\mathcal{S}$ given in the form of an infinite sum (3.20-3.21) or equivalently in the form (3.41)-(3.44)

$$\Delta\mathcal{S} = \int_{-\infty}^{\infty} \widetilde{K}(k, k) dk, \quad (5.4)$$

where

$$\widetilde{K}(k, k') = \sum_{n=1}^{\infty} \frac{1}{n} \widetilde{K}^n(k, k') \quad (5.5)$$

with

$$\begin{aligned} \widetilde{K}^n(k, k') = & \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dk_1 \cdots dk_{n-1} \times \\ & \sqrt{\frac{1}{1 + e^{\beta\epsilon(k)}}} K(k, k_1) \frac{1}{1 + e^{\beta\epsilon(k_1)}} K(k_1, k_2) \frac{1}{1 + e^{\beta\epsilon(k_2)}} \cdots \frac{1}{1 + e^{\beta\epsilon(k_{n-1})}} K(k_{n-1}, k') \sqrt{\frac{1}{1 + e^{\beta\epsilon(k')}}} \end{aligned} \quad (5.6)$$

and the dressed energy $\epsilon(k)$ entering this formula being the usual one given by the equation (3.6)

$$\epsilon(k) = e(k) - T \int_{-\infty}^{\infty} K(k', k) \ln(1 + e^{-\beta\epsilon(k')}) dk'. \quad (5.7)$$

We have shown also, that $\Delta\mathcal{S}$ and its $T \rightarrow 0$ limit exist and are finite. In the proof it is used that the particle density is a finite value fixed by the chemical potential μ . The "entropy" $\lim_{T \rightarrow 0} \Delta\mathcal{S}$ is the residual entropy not equal to the entropy of the (unique) ground state what would be zero. This is a consequence of the fact that the $T \rightarrow 0$ limit is taken after the $L \rightarrow \infty$ limit, and these limits do not commute.

We also calculated the contribution of the saddle point fluctuations to the free energy of open end systems. We have found that it is slightly different in structure (3.44). It is $-T\overline{\Delta\mathcal{S}}$ with

$$\overline{\Delta\mathcal{S}} = \int_0^{\infty} \widetilde{K}(k, k) dk + \int_0^{\infty} \widetilde{K}(k, -k) dk. \quad (5.8)$$

In field theory systems of massive relativistic particles are important in which the particle number is not regulated by a chemical potential. A system of this type is described by the Lee-Yang model [13] in which $e(k) = M_{\text{ch}}(k)$ and the kernel K is negative. This system behaves somewhat differently than the δ gas. There are strong arguments supporting the claim, that for PBC no $O(1)$ corrections should be present. Contrary to this our method

gives a $\Delta\mathcal{S}$ of the type (5.4) for this case. For large T $\exp\{-\beta\epsilon(k)\}$ is a constant for $e(k) < T$ and is zero above. In this case $\Delta\mathcal{S} \propto -\ln(T/M)$. If $T \rightarrow 0$, then $\epsilon(k) \rightarrow e(k)$, $\exp\{\beta\epsilon(k)\}$ diverges and $\Delta\mathcal{S}$ disappears. For the open end case we get a correction of the type (5.8) what also behaves differently as expected: due to the first term it diverges in the $T \rightarrow \infty$ limit. We note, that in a recent work [14] based on a different type of calculation it has been proposed, that the $O(1)$ correction in the open end case sholud be of the type of the second term in (5.8).

One has to note here the following. It is possible to define some field theoretical models as certain limits of lattice models, for example the scaling limits of the Hubbard or the Heisenberg models are closely related to the $SU(2)$ chiral invariant Gross-Neveu model [15–17]. Calculating the $O(1)$ corrections in the lattice models and taking the scaling limit afterward we expect $\lim_{T \rightarrow 0} \Delta\mathcal{S}$ involve the residual entropy of the vacuum too.

We also give a generalization of our result to other Bethe Ansatz systems. In a large class of models the densities satisfy equations of the type (4.1)

$$\rho^{(n)}(k) + \rho_h^{(n)}(k) = \sigma_n(k) + \sum_m \int K_{n,m}(k, k') \rho^{(m)}(k') dk' . \quad (5.9)$$

For these systems we find, that the contribution of the saddle point fluctuations to the free energy is $-T\Delta\mathbf{S}$ (4.7) with

$$\Delta\mathbf{S} = \sum_n \frac{1}{n} K^n , \quad (5.10)$$

where now

$$K^n = \sum_{m_1} \cdots \sum_{m_n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dk_1 \cdots dk_n \times \quad (5.11) \\ \frac{1}{1 + e^{\beta\epsilon_{m_1}(k_1)}} K_{m_1 m_2}(k_1, k_2) \frac{1}{1 + e^{\beta\epsilon_{m_2}(k_2)}} \cdots \frac{1}{1 + e^{\beta\epsilon_{m_n}(k_n)}} K_{m_n m_1}(k_n, k_1) .$$

In this formula the energies $\epsilon_n(k)$ are connected to the $e_n(k)$ bare ones by the equations

$$\epsilon_n(k) = e_n(k) - T \sum_m \int \ln(1 + e^{-\beta\epsilon_m(k')}) K_{m,n}(k', k) dk' . \quad (5.12)$$

As this formula is a general one, its convergence should be checked in any special case. We also point out, that in the special cases additional problems requiring further considerations may arise, as we illustrate on the example of the XXZ Heisenberg chain.

ACKNOWLEDGMENTS

I am grateful to Prof. L. Palla and Drs. Z. Bajnok and G. Takács for the valuable discussions. Work has been supported by OTKA under grant Nr. T 043330

APPENDIX A

In this Appendix we want to examine the approximation (2.20) reading as

$$\sum_{\{k_j\}} \exp \left\{ -\beta \sum_{k_j} (k_j^2 - \mu) \right\} \longrightarrow \exp \left\{ -\beta (\bar{k}^2 - \mu) L\rho(k) \Delta k + \ln \omega(\rho(k)) \right\} \quad (\text{A1})$$

in more detail. Suppose that \bar{k} is the mean value of the ks in the Δk interval, and that the Δk is small enough, so we may linearize around \bar{k} . Thus to get the left hand side of the above formula we have to calculate

$$\exp \left\{ -\beta (\bar{k}^2 - \mu) L\rho(k) \Delta k \right\} \sum_{\{n_i\}} \exp \left\{ 2q \sum_{i=1}^m n_i \right\}, \quad (\text{A2})$$

where $q = -\beta \bar{k} \Delta k / N$, the numbers n_i are integers or half-integers ($n_i = (N-1)/2 \pmod{1}$) satisfying $-(N-1)/2 \leq n_1 < \dots < n_i < n_{i+1} < \dots < n_m \leq (N-1)/2$ with $N = L(\rho(k) + \rho_0(k)) \Delta k$ and $m = L\rho(k) \Delta k$, and the $\sum_{\{n_i\}}$ extends over all possible n_i sets. This gives

$$\exp \left\{ -\beta (\bar{k}^2 - \mu) L\rho(k) \Delta k \right\} \prod_{i=1}^m \frac{\text{sh}((N-i+1)q)}{\text{sh}(iq)}, \quad (\text{A3})$$

what for small enough Δk yields

$$\exp \left\{ -\beta (\bar{k}^2 - \mu) L\rho(k) \Delta k + \ln \omega(\rho(k)) + \frac{q^2}{6} \sum_{i=1}^m ((N-i+1)^2 - i^2) \right\}. \quad (\text{A4})$$

Evaluating the sum and inserting the value of N and m we arrive at

$$\exp \left\{ -\beta \left(\bar{k}^2 - \mu - \frac{\beta \bar{k}^2 (\Delta k)^2}{6} \frac{\rho_h(k)}{\rho(k) + \rho_h(k)} \right) L\rho(k) \Delta k + \ln \omega(\rho(k)) \right\}, \quad (\text{A5})$$

It is not hard to see, that taking into account the quadratic nature of the spectrum would lead also to corrections not larger than $O((\Delta k)^2)$, i.e. the correction to the free energy density neglected in (2.20) ((A1)) is indeed small enough to disappear in the $\sum(\dots) \Delta k \longrightarrow \int(\dots) dk$ limit.

APPENDIX B

Here we show, that the parity prescription for the parameters J_i ($J_i = (N+1)/2 \pmod{1}$) does not destroy the accuracy of the free energy functional. Let us consider two systems, one described by the system of equations

$$Lk_i = 2\pi J_i - \sum_j^N 2 \tan^{-1} \left(\frac{k_i - k_j}{c} \right) \quad (B1)$$

with J_i being integer, the other with J_i being half-odd-integer independently from the particle number. Let us distinguish between the parameters of the two systems by primes (k'_i , J'_i) and double primes (k''_i , J''_i). (The real system is between this two: a primed solution is to be taken if the number of particles is odd, and a doubly primed one applies for an even number of particles.) Obviously the two kinds of solutions are in one-to-one correspondence: we consider a primed and a doubly primed solution one pair, if $J''_i = J'_i + \frac{1}{2}$ for all i . Due to the 'Galilei invariance' of (B1) the wavenumbers of the pairs are closely related: $k''_i = k'_i + \pi/L$. For this we can describe the pairs by the same $\rho(k)$ and $\rho_h(k)$ if the $\Delta k'$ and $\Delta k''$ intervals are the same just shifted by π/L relative to each other. Let us denote the free energy associated to a $\rho(k)$ in the two systems by $F'[\rho(k)]$ and $F''[\rho(k)]$, respectively. Their difference is

$$\Delta F[\rho(k)] = F''[\rho(k)] - F'[\rho(k)] \simeq \sum_k 2\bar{k}\pi\rho(k)\Delta k, \quad (B2)$$

where on the right hand site we dropped the primes. For a general $\rho(k)$ this can be of $O(1)$, but for those densities which contribute to Z it is much smaller: near to the equilibrium $\rho(k) = \rho_0(k) + r(k)$, and as ρ_0 is an even function of k ,

$$\Delta F[\rho(k)] \simeq \sum_{k>0} 2\bar{k}\pi (r(k) - r(-k)) \Delta k. \quad (B3)$$

As in the functional integral the major contribution comes from the region $r \propto 1/\sqrt{L}$, for densities important in calculating the saddle point contributions $\Delta F[\rho(k)] \propto 1/\sqrt{L}$ is a good estimation. This shows, that the difference in the free energies of the primed and the doubly primed systems disappears as $L \rightarrow \infty$. As, however, (B3) is also an estimate for the error made if the parity of the numbers J_i is not chosen properly we may conclude that the prescription for the J_i s ($J_i = (N-1)/2 \pmod{1}$) does not influence the $O(1)$ corrections.

APPENDIX C

In this Appendix we show, that all the eigenvalues κ of the matrix \mathbf{K} of (3.18) in the case of the δ Bose gas have a modulus less than unity. To do this we use the formula

$$\max \ln |\kappa| = \lim_{n \rightarrow \infty} \frac{1}{n} \ln |\text{Tr} \mathbf{K}^n|. \quad (C1)$$

As $\text{Tr} \mathbf{K}^n$ is definitely positive for all n the eigenvalue of largest modulus is positive, i.e.

$$\max \ln |\kappa| = \ln \kappa_{\max} \quad (\text{C2})$$

As we apply our formulas after the $\sum_k \Delta k \rightarrow \int dk$ limit is taken, we may use $\text{Tr} \mathbf{K}^n = \mathcal{K}^n$, that is

$$\ln \kappa_{\max} = \lim_{n \rightarrow \infty} \frac{1}{n} \ln \mathcal{K}^n. \quad (\text{C3})$$

For $T > 0$, using the relations

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dk_1 \cdots dk_{(n-1)} K(k, k_1) \cdots K(k_{(n-1)}, k') = \frac{1}{2\pi} \frac{2nc}{(nc)^2 + (k - k')^2}, \quad (\text{C4})$$

and

$$\int \rho_0(k) = N/L, \quad (\text{C5})$$

\mathcal{K}^n is overestimated by the formula

$$\mathcal{K}^n \leq \left(\frac{\rho_0(k)}{\rho_0(k) + \rho_{h,0}(k)} \right)_{\max}^{n-1} \frac{1}{\pi nc} \frac{N/L}{(\rho_0(k) + \rho_{h,0}(k))_{\min}} \quad (\text{C6})$$

This yields

$$\kappa_{\max} \leq \left(\frac{\rho_0(k)}{\rho_0(k) + \rho_{h,0}(k)} \right)_{\max}, \quad (\text{C7})$$

what for $T > 0$ is indeed less than one. This proves the applicability of the formulas (3.18) and the convergence of the series $\sum_n \frac{1}{n} \mathcal{K}^n$ (and also proves, that the Neumann series of Eq.(3.7) converges).

By a slight modification of the above estimations one can also prove, that the

$$\lim_{T \rightarrow 0} \sum_n \frac{1}{n} \mathcal{K}^n \quad (\text{C8})$$

also exists. For $T = 0$

$$\frac{\rho_0(k)}{\rho_0(k) + \rho_{h,0}(k)} = \begin{cases} 1 & \text{if } |k| \leq k_F; \\ 0 & \text{otherwise,} \end{cases} \quad (\text{C9})$$

where k_F is a finite wavenumber, under which all, above which none of the states are filled. Observing, that

$$\begin{aligned} \int_{-k_F}^{k_F} \cdots \int_{-k_F}^{k_F} dk_1 \cdots dk_{(n-1)} K(k, k_1) \cdots K(k_{(n-1)}, k) &\leq \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dk_1 \cdots dk_{(n-1)} K(k, k_1) \cdots K(k_{(n-1)}, k) &= \frac{1}{\pi n c} \end{aligned} \quad (\text{C10})$$

we see, that

$$\lim_{T \rightarrow 0} \sum_n \frac{1}{n} \mathcal{K}^n \leq \frac{2k_F}{\pi c} \sum_n \frac{1}{n^2} = \frac{2k_F \pi}{6c}. \quad (\text{C11})$$

We note, that if $c \rightarrow 0$, i.e. if the kernel is a δ -function, both for $T > 0$ and $T \rightarrow 0$ our estimation blows up.

APPENDIX D

In this appendix we discuss the questions connected to the cutoff procedure involved in the evaluation of the free energy. First we notice, that this problem is rather a problem of the accuracy of the macroscopic part of the free energy. To see this consider the partition function Z of a system. The macroscopic part of the free energy is defined as

$$F_{\text{mac}} = Lf, \quad \text{where} \quad f = -T \lim_{L \rightarrow \infty} \frac{1}{L} \ln Z. \quad (\text{D1})$$

The next to leading correction to the macroscopic part is $O(L^0)$ if

$$(\Delta \mathcal{S} =) \lim_{L \rightarrow \infty} \ln (Z e^{\beta F}) \quad (\text{D2})$$

is finite but zero. If so, the free energy defined through the logarithm of the partition function behaves for large enough L as

$$F = -T \ln Z = Lf - T\Delta \mathcal{S}. \quad (\text{D3})$$

In our case in order to apply Stirling's formula we have to introduce the cutoff Λ in the k space, and for the same reason we have to calculate sums instead of integrals. This way the partition function we obtain is of the form

$$Z_L(\Lambda, \Delta k) = e^{-\beta F_{\text{min}}(L, \Lambda, \Delta k) + \Delta S(\Lambda, \Delta k)} = e^{-\beta L f_{\text{min}}(\Lambda, \Delta k) + \Delta S(\Lambda, \Delta k)}. \quad (\text{D4})$$

To make our reasoning simpler, for the time being we suppose, that it is accurate enough to replace the summations on Δk by integrals. This way we have

$$Z_L(\Lambda) = e^{-\beta L f_{\min}(\Lambda) + \Delta S(\Lambda)}. \quad (\text{D5})$$

Now

$$-T \lim_{L \rightarrow \infty} \frac{1}{L} \ln Z_L = f_{\min}(\Lambda) \quad (\text{D6})$$

from which the free energy density f is obtained through a next limit

$$f = \lim_{\Lambda \rightarrow \infty} f_{\min}(\Lambda). \quad (\text{D7})$$

This leads to difficulties in filtering out the next to leading corrections, as

$$\lim_{L \rightarrow \infty} \ln (Z_L(\Lambda) e^{\beta L f}) \rightarrow (+\text{or}-)\infty \quad \text{like} \quad L(f - f_{\min}(\Lambda)), \quad (\text{D8})$$

i.e. $f - f_{\min}(\Lambda)$ *hides* the correction we want to get. (We have to emphasize, that the value of the saddle point corrections is not effected by this, nevertheless we have to see the leading order term more accurately than the correction we expect.)

To resolve this problem we propose the following. Taking larger and larger L allows taking larger and larger Λ , thus the two limits can be synchronized: a $\Lambda(L)$ can be chosen so, that

$$\Lambda(L) \xrightarrow[L \rightarrow \infty]{} \infty, \quad (\text{D9})$$

while the condition for applying Stirling's formula within the cutoffs is met, i.e.

$$L \Delta k \rho_0(\Lambda) \gg 1. \quad (\text{D10})$$

It seems plausible, that if it is possible

- to choose $\Lambda(L)$ so, that $L(f - f_{\min}(\Lambda(L))) \rightarrow 0$,
- and take also $\Delta k \rightarrow 0$ so that the difference between the sums and integrals disappears,

while (D10) holds, than the next to leading order correction to the free energy is $\lim_{\Lambda \rightarrow \infty} \Delta S(\Lambda)$ indeed.

In the following we argue, that for the repulsive δ Bose gas one can construct an appropriate cutoff procedure. First let us consider (D10). Due to the BA equations

$$\rho_0(\Lambda) = \sigma \frac{e^{-\beta \epsilon(\Lambda)}}{1 + e^{-\beta \epsilon(\Lambda)}} (1 + O(K(\Lambda, 0))) \sim \sigma e^{-\beta \epsilon(\Lambda)} \sim \sigma e^{-\beta e(\Lambda)}, \quad (\text{D11})$$

thus we require

$$L\Delta k\sigma e^{-\beta e(\Lambda)} \rightarrow \infty. \quad (\text{D12})$$

(Here we used also, that for large Λ $\epsilon(\Lambda) = e(\Lambda) + O(K(\Lambda, 0))$.)

The effect of introducing integrals instead of the sums can be estimated by an Euler-Maclaurin type formula. We find the most significant part is

$$\sim L(\Delta k)^2 \frac{d}{d\Lambda} \sigma e^{-\beta \epsilon(\Lambda)}, \quad (\text{D13})$$

i.e. we need

$$L(\Delta k)^2 \sigma e^{-\beta e(\Lambda)} \Lambda \rightarrow 0. \quad (\text{D14})$$

(We do not give details here, just note, that making the error due to (2.20) (Appendix A) to disappear fast enough leads to the condition (D14) too.)

Finally we had to estimate $L(f - f_{\min}(\Lambda(L)))$ but for this we have to specify the procedure. One possibility is simply to omit all the modes outside the $\pm\Lambda$ interval (corresponding to taking their energy equal to ∞), but in this scheme $f_{\min}(\Lambda(L))$ does not converge in Λ to f fast enough. A procedure providing a much faster convergence can be constructed realizing that the particles of high energy behave as free ones. In this scheme the free energy of the system is built up of two parts: the contribution of the modes within the $\pm\Lambda$ interval is calculated using Stirling's formula (just as in the bulk of the paper), while the contribution of the modes outside the cutoffs is approximated by the contribution of free particles of energy $e(k)$ with density of states $\rho(k) + \rho_h(k)$ given by the BA equations. In this scheme the minimization of the free energy leads to a dressed energy given by the equation

$$\begin{aligned} \epsilon_{\Lambda}(k) = & e(k) - T \int_{-\Lambda}^{\Lambda} K(k', k) \ln \left(1 + e^{-\beta \epsilon_{\Lambda}(k')} \right) dk' \\ & - T \left(\int_{-\infty}^{-\Lambda} + \int_{\Lambda}^{\infty} \right) K(k', k) \ln \left(1 + e^{-\beta e(k')} \right) dk', \end{aligned} \quad (\text{D15})$$

and the free energy density is

$$f_{\min}(\Lambda) = - \int_{-\Lambda}^{\Lambda} \ln \left(1 + e^{-\beta \epsilon_{\Lambda}(k)} \right) \sigma dk - \left(\int_{-\infty}^{-\Lambda} + \int_{\Lambda}^{\infty} \right) \ln \left(1 + e^{-\beta e(k)} \right) \sigma dk. \quad (\text{D16})$$

The leading part of $L(f - f_{\min}(\Lambda(L)))$ is of the order of

$$\sim L \int_{\Lambda}^{\infty} \left(\ln \left(1 + e^{-\beta e(k)} \right) - \ln \left(1 + e^{-\beta \epsilon(k)} \right) \right) \sigma dk, \quad (\text{D17})$$

what through some straightforward manipulations and approximations leads to the condition

$$Le^{-\beta e(\Lambda)} \frac{\sigma}{\Lambda^3} \rightarrow 0. \quad (\text{D18})$$

It is not hard to see, that it is possible to define an $L \rightarrow \infty$, $\Lambda \rightarrow \infty$ and $\Delta k \rightarrow 0$ limit so, that all the conditions (D12)(D14) and (D18) are satisfied.

We have to note, that the above reasoning concerning the existence of appropriate cutoff procedure works for the δ Bose gas only, for other systems it has to be reformulated, but for certain models it is also possible, that there is no need for such a cutoff procedure.

APPENDIX E

In the present Appendix we discuss a case of the δ Bose gas with open ends, in which surface bound states can be present. The system is described by the Bethe Ansatz equations

$$2Lk_j = 2\pi J_j - \varphi_0(k_j) - \varphi_L(k_j) - \sum_{l(\neq j)}^N \left\{ 2 \tan^{-1} \left(\frac{k_j - k_l}{c} \right) + 2 \tan^{-1} \left(\frac{k_j + k_l}{c} \right) \right\}. \quad (\text{E1})$$

Now we suppose, $\varphi_L = \pi$ corresponding to an infinitely high wall closing the chain at L , but we take

$$\varphi_0(k) = \pi - \tan^{-1} \frac{k}{\gamma}. \quad (\text{E2})$$

Also this corresponds to an infinitely high wall, but this wall is preceded by an infinitely deep, but also infinitely narrow potential well. An appropriate tuning of the width and depth of the well leads to the above reflection phase shift. (The effect of such a potential in case of a δ Fermi gas is discussed in [7], and a similar case of a Hubbard chain in [8]) This potential can always generate at least one surface bound state: it is not hard to see, that for any distribution of the real ks (we denote them by Latin ks) (E1) has also imaginary solution too corresponding to a surface bound state at the end at $x = 0$. Denoting this by $i\kappa$ we find

$$\kappa = \gamma - \delta, \quad \delta = 2\gamma e^{-2L\gamma} \prod_i \frac{(c - \gamma)^2 + k_i^2}{(c + \gamma)^2 + k_i^2} \quad (\text{E3})$$

In the presence of such an imaginary wavenumber the real ones satisfy the equations

$$2Lk_j = 2\pi J_j + \tan^{-1} \frac{k_j}{\gamma} - t(k_j, \kappa) - \sum_{l(\neq j)}^N \left\{ 2 \tan^{-1} \left(\frac{k_j - k_l}{c} \right) + 2 \tan^{-1} \left(\frac{k_j + k_l}{c} \right) \right\}, \quad (\text{E4})$$

with

$$t(k, \kappa) = \left\{ 2 \tan^{-1} \left(\frac{k - i\kappa}{c} \right) + 2 \tan^{-1} \left(\frac{k + i\kappa}{c} \right) \right\}. \quad (\text{E5})$$

If the potential is strong enough, i.e. $\gamma > c$ (E1) has also solutions with more than one imaginary wavenumbers. These describe more than one particles bound to the $x = 0$ end. For two imaginary wavenumbers $i\kappa_{1,2}$ we find

$$\begin{aligned} \kappa_1 &= \gamma - \delta_1, & \delta_1 &\simeq 2\gamma \left(\frac{\gamma - c}{\gamma} \right)^2 e^{-2L(2\gamma - c)} \prod_i \frac{(\gamma - c)^2 + k_i^2}{(\gamma + c)^2 + k_i^2} \frac{(\gamma - 2c)^2 + k_i^2}{(\gamma)^2 + k_i^2} \\ \kappa_2 &= \kappa_1 - c - \delta_2, & \delta_2 &\simeq 2c \left(\frac{\gamma - c}{\gamma} \right) e^{-2L(\gamma - c)} \prod_i \frac{(\gamma - 2c)^2 + k_i^2}{(\gamma)^2 + k_i^2}. \end{aligned} \quad (\text{E6})$$

In general, a solution with ν imaginary wavenumbers $i\kappa_\alpha$, $\alpha = 1, \dots, \nu$ exists, if $\gamma - (\nu - 1)c > 0$. The $i\kappa_\alpha$ s have the form

$$\kappa_1 = \gamma - \delta_1, \quad \kappa_\alpha = \kappa_{\alpha-1} - \delta_\alpha \quad (\alpha = 2, \dots, \nu) \quad (\text{E7})$$

with all δ_α being exponentially small in L . In these solutions the real k set satisfies the equations

$$2Lk_j = 2\pi J_j + \tan^{-1} \frac{k_j}{\gamma} - t_\nu(k_j) - \sum_{l(\neq j)}^N \left\{ 2 \tan^{-1} \left(\frac{k_j - k_l}{c} \right) + 2 \tan^{-1} \left(\frac{k_j + k_l}{c} \right) \right\}, \quad (\text{E8})$$

with

$$t_\nu(k) = \sum_{\alpha=1}^{\nu} t(k, \kappa_\alpha). \quad (\text{E9})$$

The thermodynamic treatment follows the procedure described in the bulk of the paper with the difference, that $\varphi_0(k) + \varphi_L(k)$ should be replaced by $-2 \tan^{-1}(k/\gamma) + t_\nu(k)$, and the calculation should be repeated for all possible ν . Now the minimal free energy at a given ν (taking also into account the direct energy contributions of the imaginary wavenumbers) is

$$\overline{F}_\nu = F_{\min} + \Delta F + \Delta F_\nu \quad (\text{E10})$$

with F_{\min} being the bulk value, ΔF given by (3.39), and

$$\Delta F_{\nu=0} = \frac{T}{2\pi} \int_0^\infty \frac{2\gamma}{\gamma^2 + k^2} \ln \left(1 + e^{-\beta\epsilon(k)} \right) dk, \quad \Delta F_{\nu>0} = \Delta F_0 + \sum_{\alpha=1}^{\nu} \epsilon(i\kappa_\alpha), \quad (\text{E11})$$

with $\epsilon(i\kappa)$ being the formal extension of (3.6) to complex k . As the saddle point contribution is independent of the state of the surface finally we arrive at

$$\bar{Z} = \sum_{\nu=0}^n e^{-\beta \bar{F}_\nu + \Delta S} = e^{-\beta(F_{\min} + \Delta F) + \Delta S} \sum_{\nu=0}^n e^{-\beta \Delta F_\nu} \quad (\text{E12})$$

with n being the maximal possible value of ν (that is always larger than or equal to one).

APPENDIX F

In this appendix we try to estimate the effect of the

$$-T \ln(2\mathcal{L} + 1) \quad (\text{F1})$$

term in the free energy of an isotropic Heisenberg chain in no magnetic field. As \mathcal{L} is the magnetization of the highest weight member of the multiplet (which is the one described by the BA equations)

$$\mathcal{L} = \frac{N}{2} - N \sum_{(n)} l_{(n)} \sum \rho^{(n)}(k) \Delta k. \quad (\text{F2})$$

For the densities $\rho_0^{(n)}$ given by the minimization of the leading part of the free energy functional this is zero, thus

$$\mathcal{L} = -N \sum_{(n)} l_{(n)} \sum r^{(n)}(k) \Delta k, \quad (\text{F3})$$

with

$$r^{(n)}(k) = \rho^{(n)}(k) - \rho_0^{(n)}(k). \quad (\text{F4})$$

In analogy with (2.18) new variables are introduced

$$\xi^{(n)}(k) = \sqrt{N \Delta k \frac{1}{2} \frac{\rho_0^{(n)}(k) + \rho_{h,0}^{(n)}(k)}{\rho_0^{(n)}(k) \rho_{h,0}^{(n)}(k)} r^{(n)}(k)} \quad (\text{F5})$$

leading to

$$\mathcal{L} = - \sum_{(n)} l_{(n)} \sum \sqrt{2N \Delta k \frac{\rho_0^{(n)}(k) \rho_{h,0}^{(n)}(k)}{\rho_0^{(n)}(k) + \rho_{h,0}^{(n)}(k)} \xi^{(n)}(k)}. \quad (\text{F6})$$

As the application of Stirling's formula is correct if

$$N\Delta k\rho_0^{(n)}(k) \gg 1 \quad \text{and} \quad N\Delta k\rho_{h,0}^{(n)}(k) \gg 1, \quad (\text{F7})$$

– just as in the case of free Fermi or δ Bose gas – one has to make sure through a cutoff procedure, that these inequalities hold. As the main contribution of the saddle point fluctuations come from the $|\xi^{(n)}(k)| \sim O(1)$ region we may conclude that for the thermodynamically important states \mathcal{L} can be large.

The contribution of the (F1) can be estimated as follows. Denoting the leading part of the free energy of the system as a function of the magnetization by $\mathbf{F}(S)$, the corrected free energy

$$\mathbf{F}(\mathcal{L}) - T \ln(2\mathcal{L} + 1) \quad (\text{F8})$$

has to be minimized. This leads to

$$\mathcal{L} \simeq \sqrt{T\chi}, \quad (\text{F9})$$

where the susceptibility

$$\chi = \left(\frac{\partial^2 \mathbf{F}(S)}{\partial S^2} \Big|_{S=0} \right)^{-1} \propto N. \quad (\text{F10})$$

This way the minimal value is

$$\mathbf{F}(0) - T \ln(2\sqrt{T\chi} + 1) \quad (\text{F11})$$

indicating that the next to leading term is $O(\ln N)$, what is much larger than the $O(1)$ contribution of the saddle point fluctuations.

REFERENCES

- [1] Ian Affleck, Andreas W.W. Ludwig, Phys.Rev.Lett. **67** 161 (1991);
 A. LeClair, G. Mussardo, H. Saleur, S. Shorik, Nucl.Phys.B **453** (1995) 581;
 P. Dorey, I. Runkel, R. Tateo, G. Watts, Nucl.Phys.B **578** (2000) 85
- [2] W.B. Thimm, J. Kroha, J. von Delft, Phys.Rev.Lett. **82** 2143 (1999);
 P. Simon, I. Affleck, Phys.Rev.Lett. **89** 206602 (2002);
- [3] C.N. Yang, C.P. Yang, J.Math.Phys. **10** (1969) 1115.
- [4] In the treatment of Yang and Yang the chemical potential is not added to the energy at the beginning, but it enters the procedure when minimizing the free energy as the Lagrange multiplier connected to the condition $N = \text{const}$. In our calculation we have to lift this condition, thus we have to use the chemical potential μ as a variable instead of the particle number N .
- [5] Handbook of Mathematical Functions, Ed.: M. Abramowitz and I.A. Stegun, Dover Publications, Inc. New York,
- [6] E. Lieb, W. Liniger, Phys.Rev. **130**, 1605 (1963)
- [7] F. Woynarovich, Phys.Lett. **108A** 401 (1985)
- [8] G. Bedürftig, H. Fram; J.Phys.A **30** (1997) 4139
- [9] M. Takahashi, M. Suzuki, Progr.Theor.Phys. **48** 2187 (1972)
- [10] M. Takahashi, Progr.Theor.Phys. **47** 69 (1972)
- [11] T. Deguchi, F.H.L. Essler, F. Gohmann, A. Klumper, V.E. Korepin, K. Kusakabe, Phys.Rep. **331** (2000) 197
- [12] A. Virosztek, F. Woynarovich, J.Phys.A **17** (1984) 3029
- [13] T.D. Lee and C.N. Yang, Phys.Rev. **87** (1952) 404; Al.B. Zamolodchikov, Nucl. Phys.B **342** (1990) 695
- [14] P. Dorey, D. Fioravanti, Chaiho Rim, R. Tateo, hep-th/0404014
- [15] F. Woynarovich, P. Forgács, Nucl.Phys.B **498** [FS] (1997) 565
- [16] F. Woynarovich, P. Forgács, Nucl.Phys.B **538** [FS] (1999) 701
- [17] T. Hauer, A. Rákos, F. Woynarovich, Nucl.Phys.B **562(3)** [FS] (1999) 497